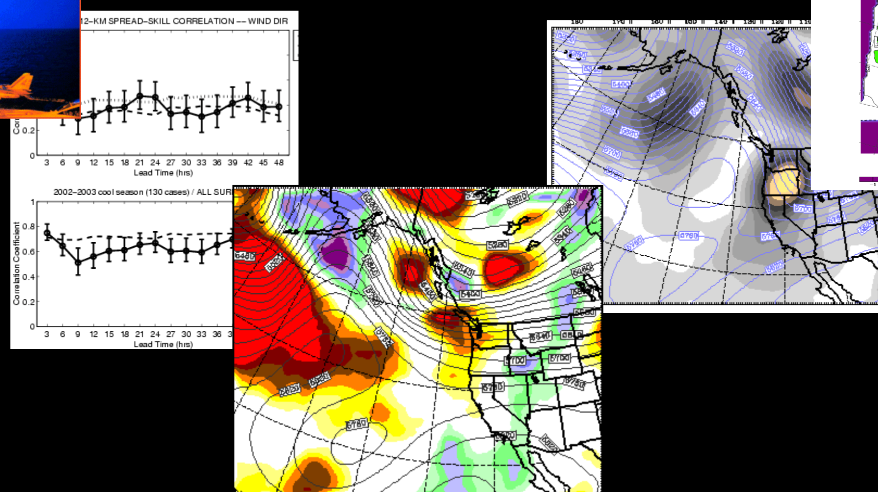
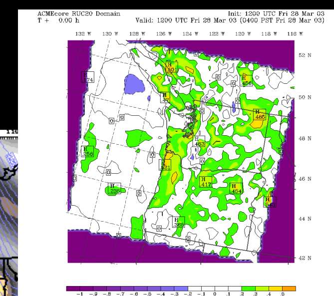
Toward Short-Range Ensemble Prediction of Mesoscale Forecast Error

Eric P. Grimit and Clifford F. Mass
University of Washington



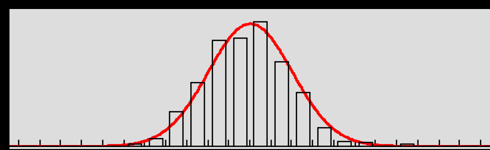


Supported by:
ONR Multi-Disciplinary University Research Initiative (MURI) and
A Consortium of Federal and Local Agencies

Traditional Approach – Spread-Error Correlation

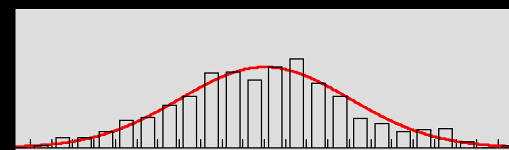
- Ensemble spread should provide an approximation to the true forecast uncertainty

agreement



better forecast reliability

disagreement



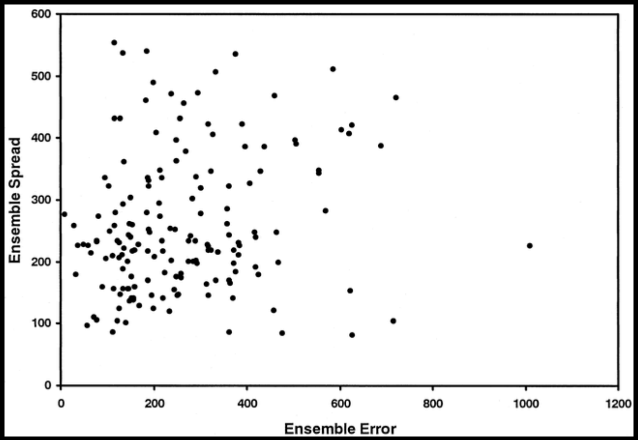
worse forecast reliability

- To quantify this "spread-skill relationship":
  - Find the linear correlation between ensemble spread (σ) and the ensemble mean forecast error ($|e_{EM}|$) over a large sample
  - Strength of the correlation is limited by the case-to-case spread variability (β) (Houtekamer, 1993; Whitaker and Loughe, 1998)

$$\rho^2(\sigma,|e_{EM}|) = \frac{2}{\pi}\;\frac{1-\exp(-\beta^2)}{1-\frac{2}{\pi}\exp(-\beta^2)}\;;\ \beta=\mathrm{std}(\ln\sigma)$$

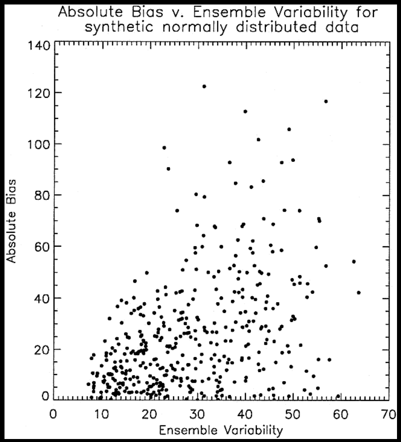
Observed Forecast Error Predictability: A Disappointment

Tropical Cyclone Tracks



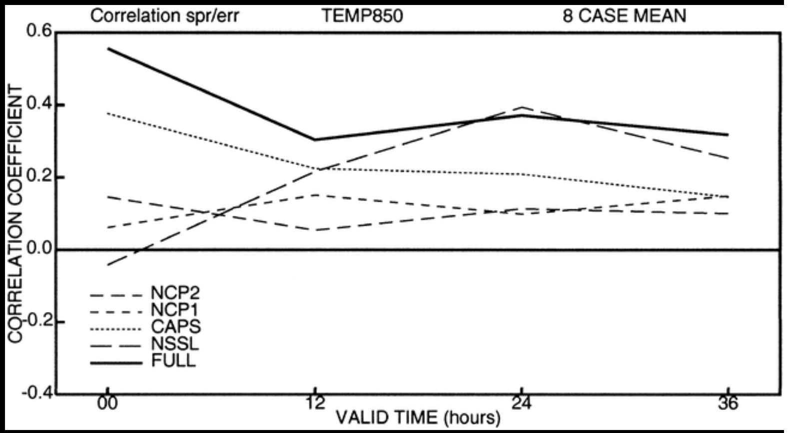
[c.f. Goerss 2000]

NCEP SREF Precipitation



[c.f. Hamill and Colucci 1998]

SAMEX '98 SREFs



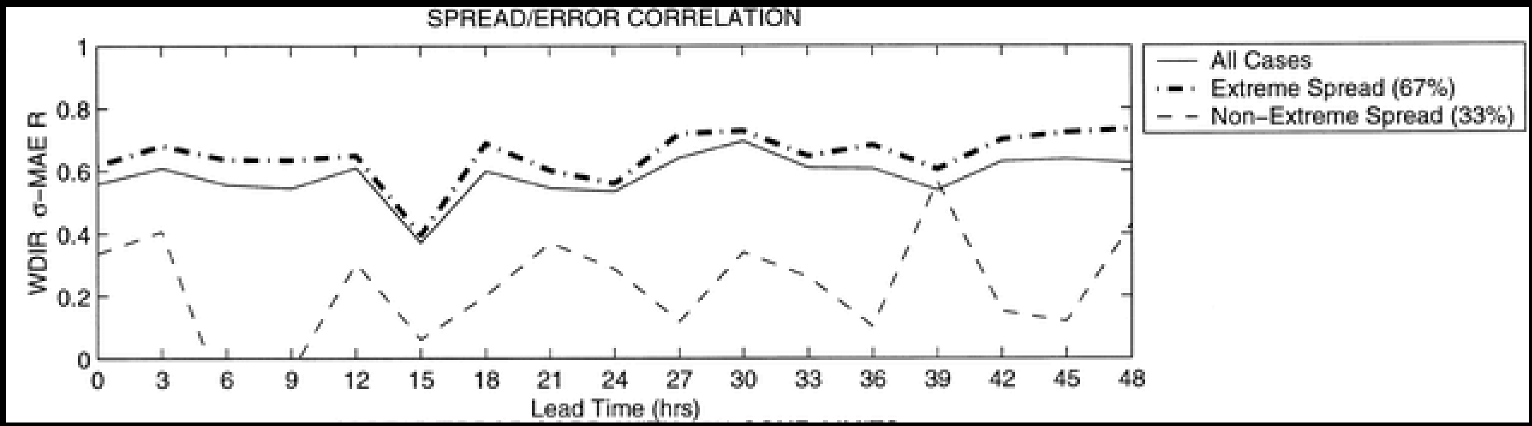
[c.f. Hou et al. 2001]

- **Highly scattered relationships, thus low correlations**
- **No indication of spread-error correlation potential**
- **No assessment of dependency on the metrics used**

Observed Forecast Error Predictability: A Disappointment

Not all hope is lost…

UW MM5 SREF 10-m Wind Direction



[c.f. Grimit and Mass 2002]

- **More recent studies show that domain-averaged spread-error correlations can be as high as 0.6-0.7**

  (Grimit and Mass 2002, Stensrud and Yussouf 2003)

  - **Potentially higher correlations can be achieved by considering only cases with extreme spread**

A Simple Stochastic Model of Spread-Skill

PURPOSES:

1) To establish <u>practical</u> limits of forecast error predictability, that could be expected given perfect ensemble forecasts of finite size.

2) To address the user-dependent nature of forecast error estimation by employing a variety of predictors and error metrics.

A Simple Stochastic Model of Spread-Skill

1. Draw today's "forecast uncertainty" from a log-normal distribution (Houtekamer 1993 model).
   $$\ln(\sigma) \sim N(\ln(\sigma_f), \beta^2)$$

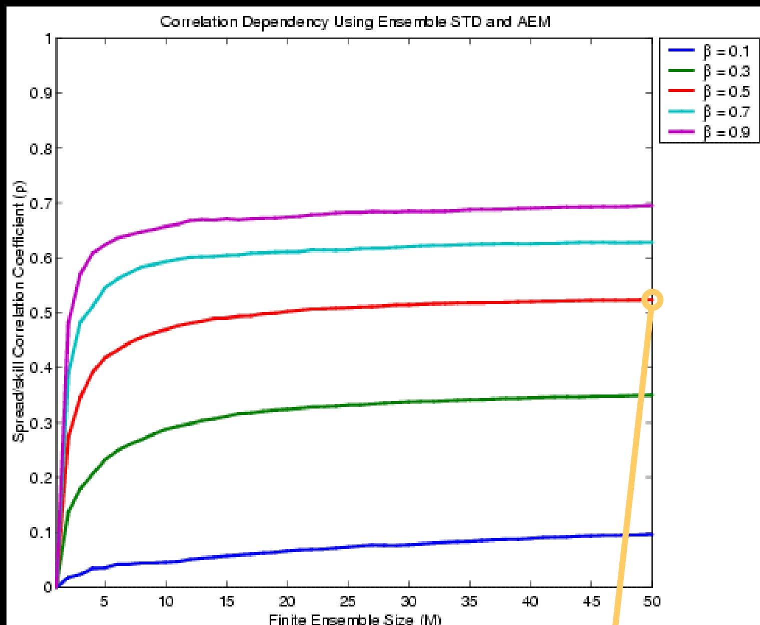
2. Create synthetic ensemble forecasts by drawing M values from the "true" distribution.
   $$F_i \sim N(Z, \sigma^2) \; ; \; i = 1,2,\ldots,M$$

3. Draw the verifying observation from the same "true" distribution (statistical consistency).
   $$V \sim N(Z, \sigma^2)$$

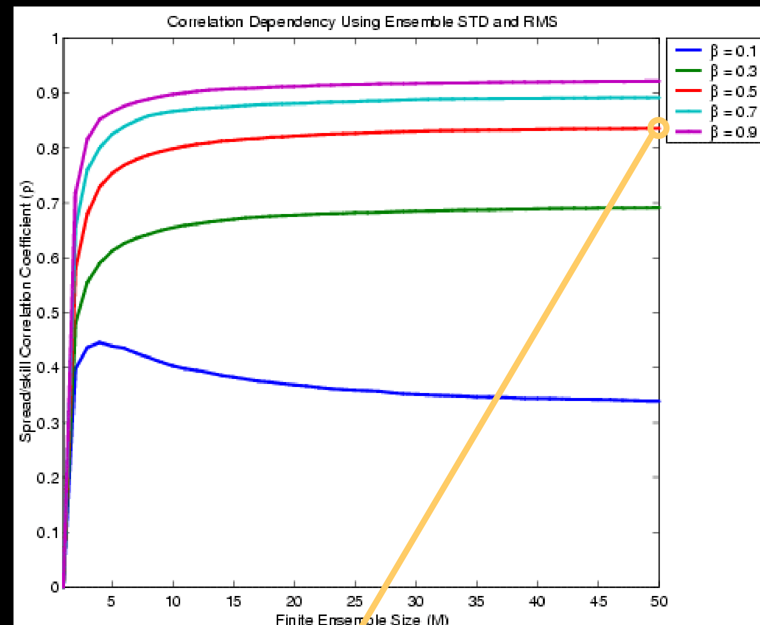
- Stochastically simulated ensemble forecasts at a single, arbitrary observing location or model-grid box with 50,000 realizations (cases)
- Assumed:
  - Gaussian statistics
  - statistically consistent (perfectly reliable) ensemble forecasts
- Varied:
  - temporal spread variability (β)
  - finite ensemble size (M)
  - spread and skill metrics (continuous and categorical)

Simple Model Spread-Error Correlations

STD-AEM correlation



STD-RMS correlation



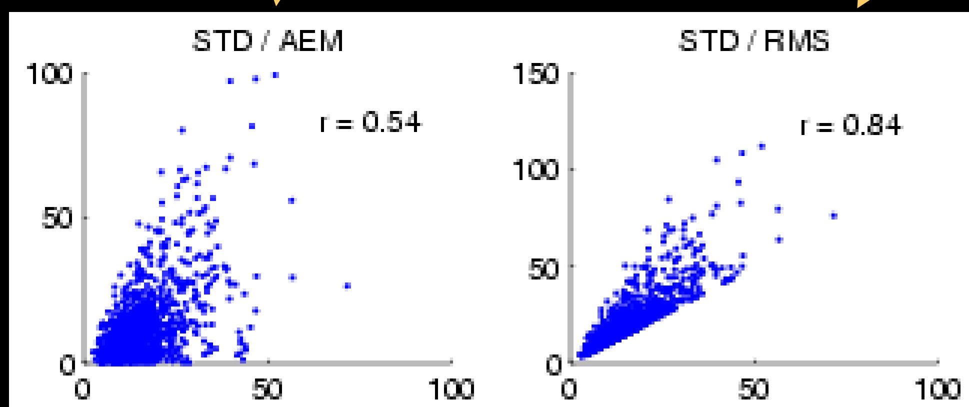
spread

STD = Standard Deviation

error

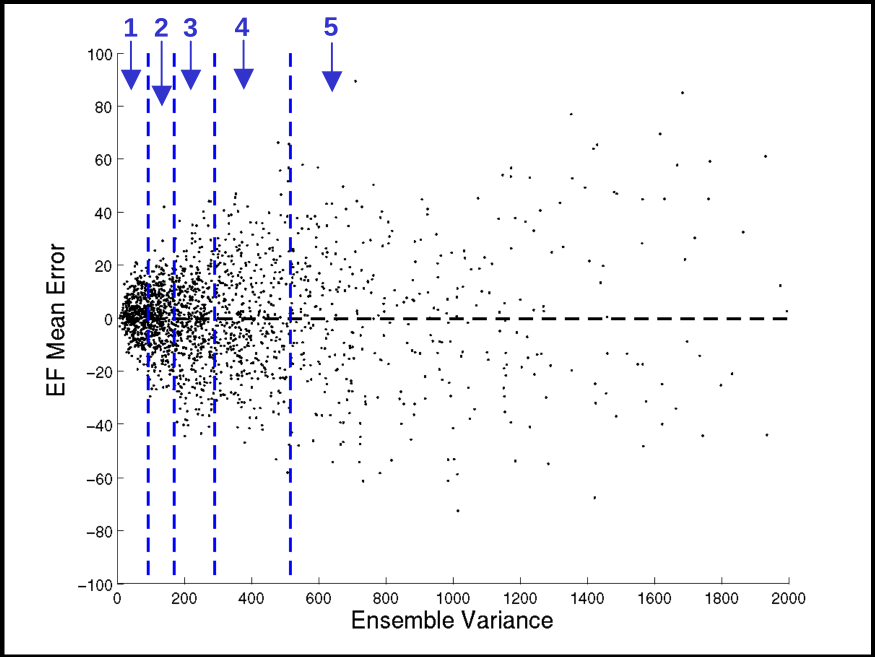
RMS= Root-Mean Square error

AEM= Absolute Error of the ensemble Mean



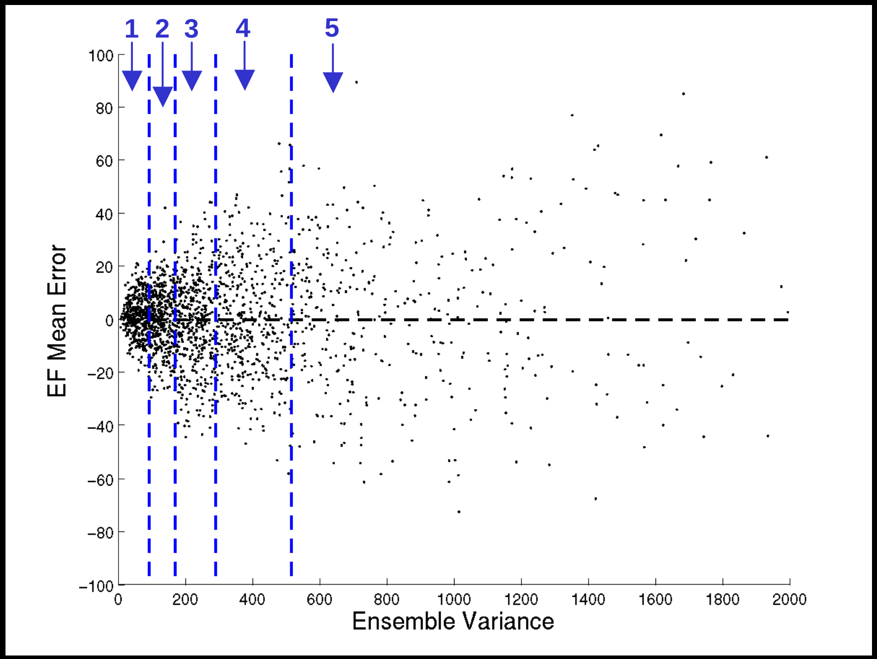
$\beta = 0.5$;
$M = 50$

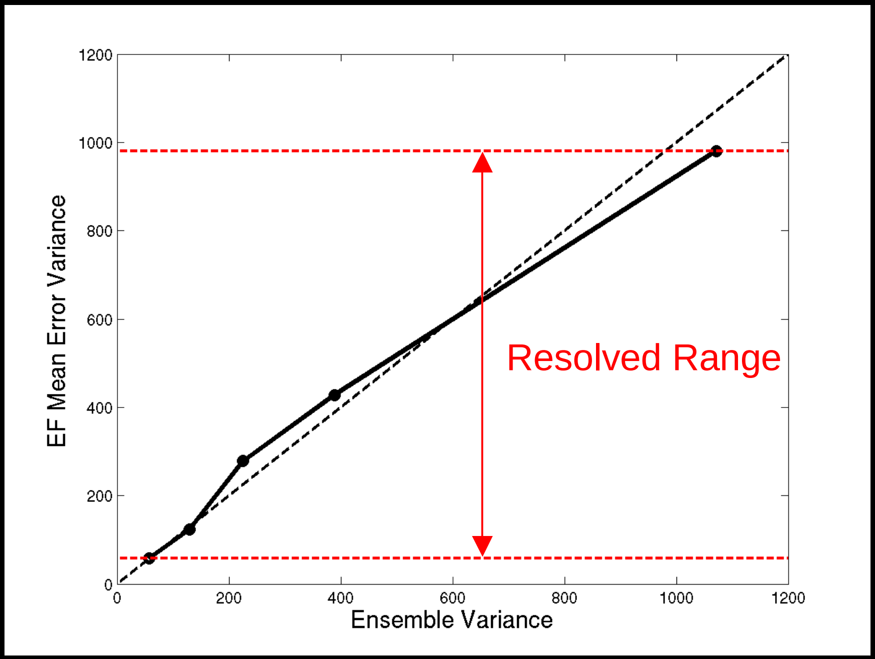
Alternative Approaches



Given statistical consistency, ensemble variance should equal the EF mean forecast error variance.

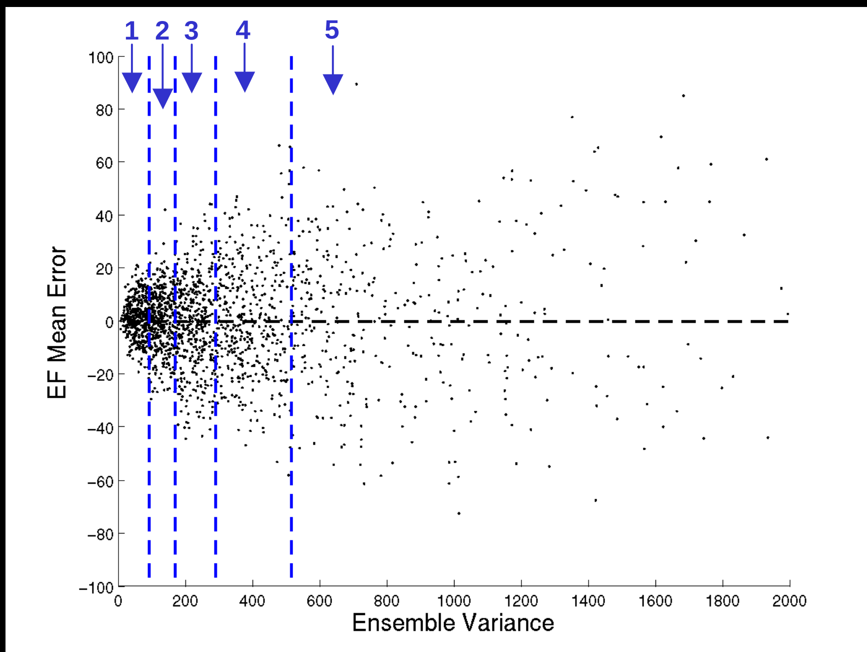
Alternative Approaches

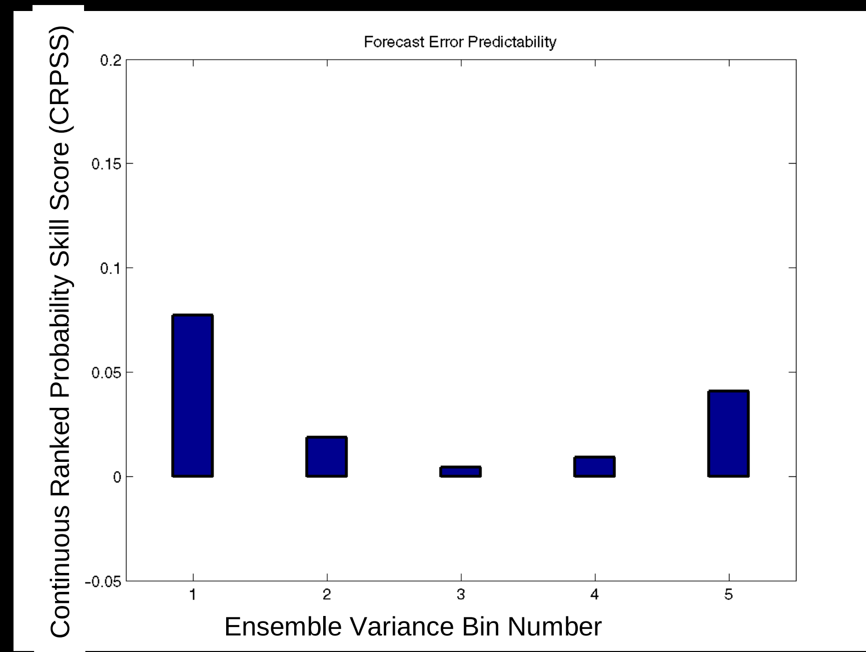




- Resolved range of error variance (Wang and Bishop 2003)
  - Choose N_{bin} equally populated bins of ensemble variance
  - Find the mean ensemble variance and the error variance within each bin
  - The range of resolved error variances indicates closeness to statistical consistency
  - Could also be applied to other error metrics (e.g. AEM, RPS)

Alternative Approaches





- Probabilistic skill of forecast error predictions
  - Use errors conditioned by spread category as probabilistic predictions of forecast error.
    - Evaluate using CRPS and its associated skill score with a cross-validation procedure.
    - CRPSS measures the continuous forecast error predictability.
  - For categorical error forecasts, use BS or RPS and the associated skill score.
  - Tradeoff between bin widths and number of samples in each bin.

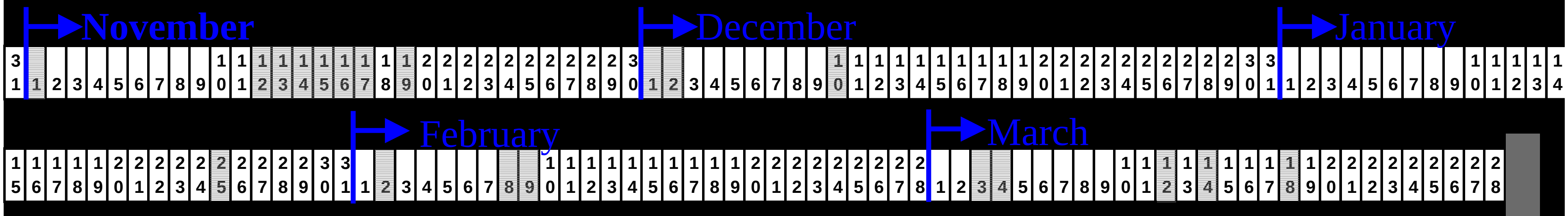
UW SREF System Summary

	Name	# of Members	EF Type	Initial Conditions	Forecast Model(s)	Forecast Cycle	Domain
Homegrown	ACME	17	SMMA	8 Ind. Analyses, 1 Centroid, 8 Mirrors	"Standard" MM5	00Z	36km, 12km
Homegrown	ACME[core]	8	SMMA	Independent Analyses	"Standard" MM5	00Z	36km, 12km
Homegrown	ACME[core+]	8	PMMA	" "	8 MM5 variations	00Z	36km, 12km
Imported	PME	8	MMMA	" "	8 "native" large-scale	00Z, 12Z	36km

ACME: Analysis-Centroid Mirroring Ensemble
PME: Poor-Man's Ensemble

SMMA: Single-Model Multi-Analysis
PMMA: Perturbed-Model Multi-Analysis
MMMA: Multi-model Multi-Analysis

Mesoscale SREF and Verification Data



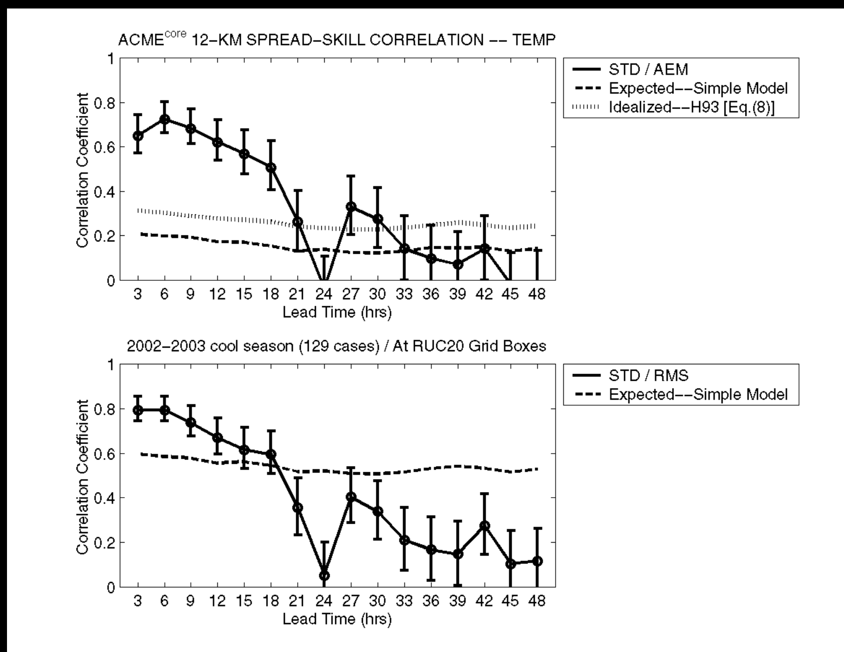
- Mesoscale SREF Data:
  - Total of 129, 48-h forecasts (31 OCT 2002 – 28 MAR 2003) all initialized at 0000 UTC
  - Missing forecast case days are shaded
  - Parameters of Focus:
    - 12 km Domain: Wind @ 10m (**$WDIR_{10}$, $WSPD_{10}$**), Temperature at 2m (**T_2**)
  - Short-term mean bias correction
    - Applied at every location and forecast lead time separately
    - Varied training window from 2-30 days

- Verification Data:
  - 12 km Domain: RUC20 analysis (NCEP 20 km mesoscale analysis) observations

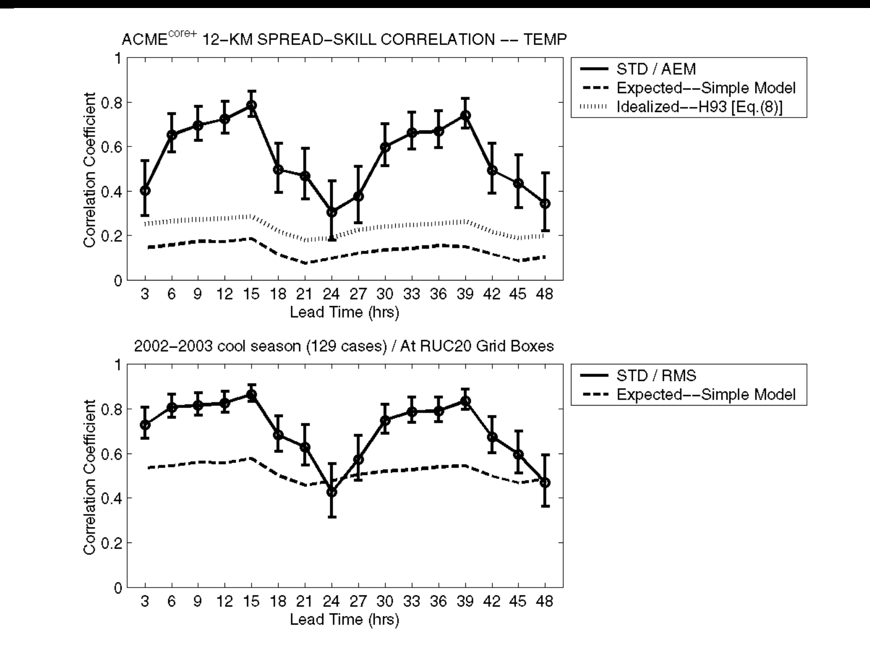
Domain-Averaged Spread-Error Correlation

(no bias correction)

ACME^{core}



ACME^{core+}

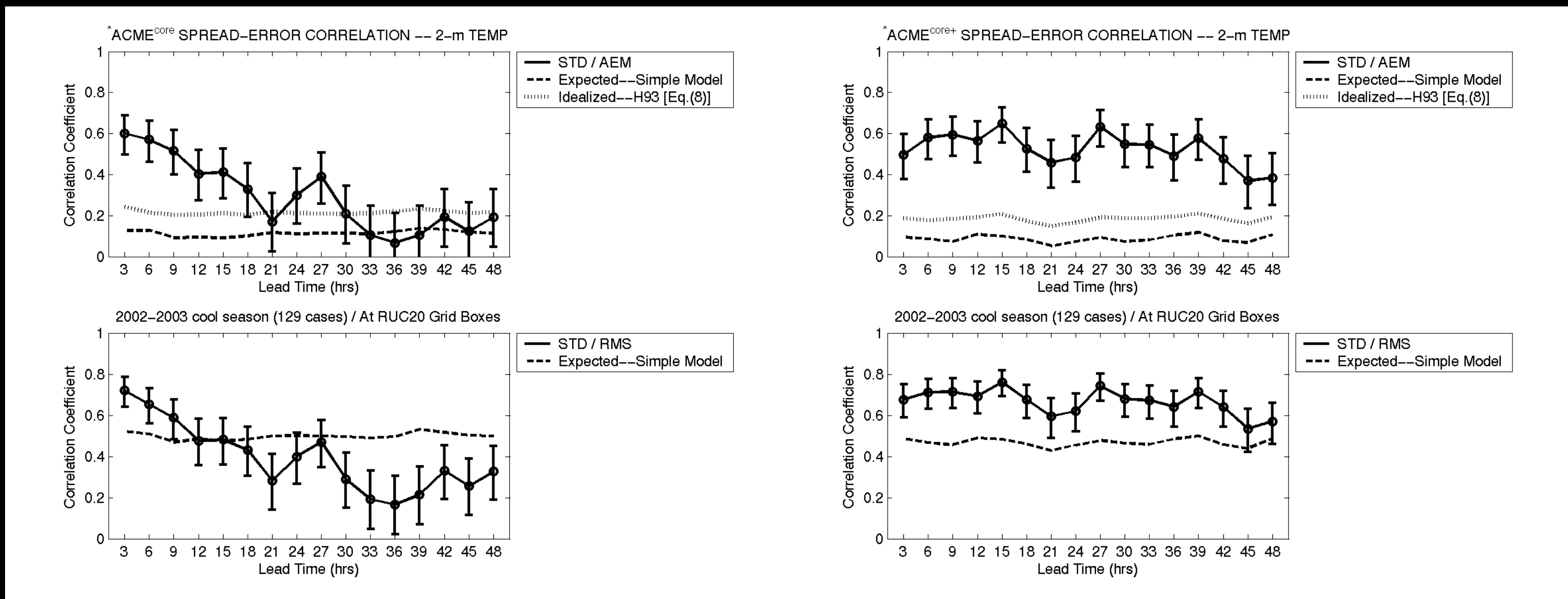


- The benefit of including model physics variability is apparent.
- Domain-averaging produces correlations much higher than expected. Correlations of averages are referred to as *ecological correlations* in statistics.

Domain-Averaged Spread-Error Correlation

(14-day bias correction)

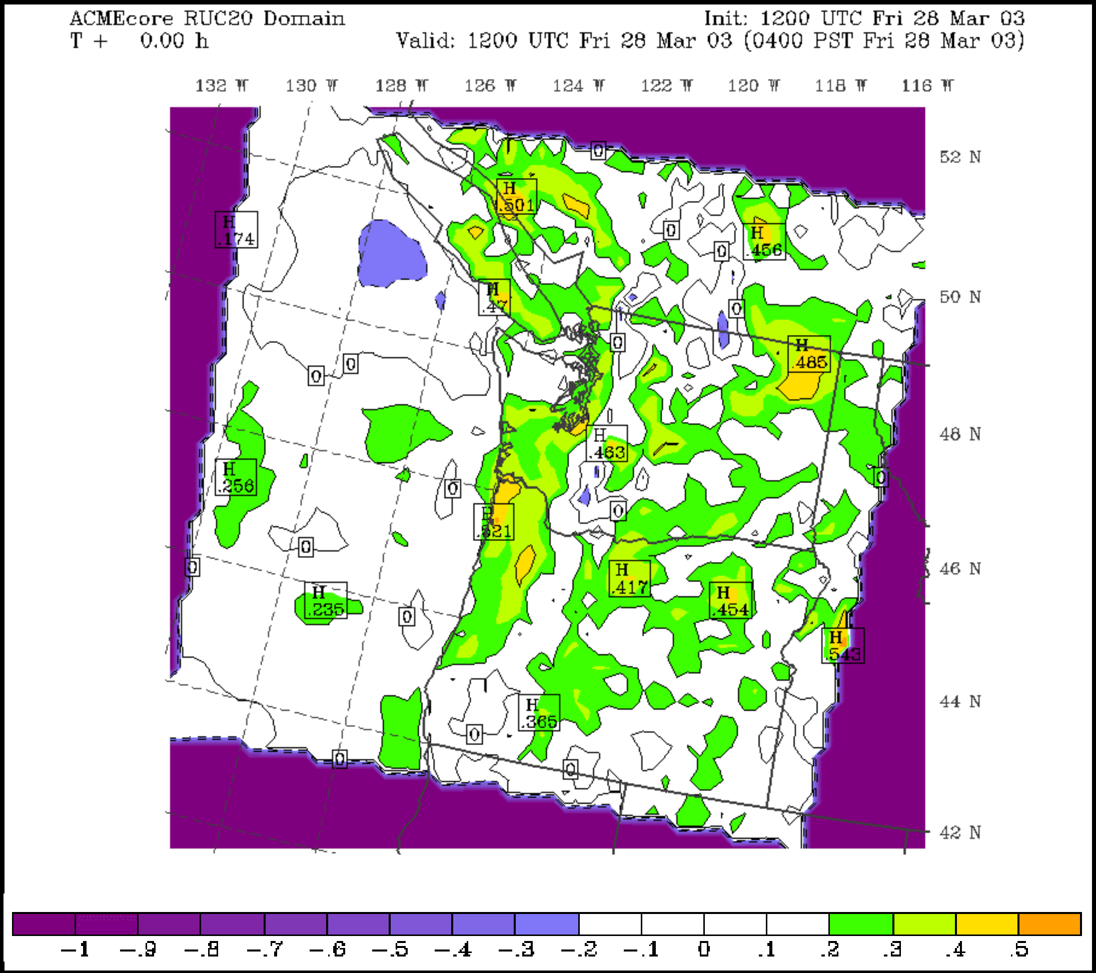
*ACME^{core} *ACME^{core+}



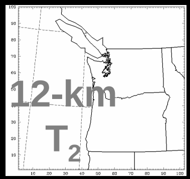
- Bias correction reduces case-to-case spread variability, resulting in poorer spread-error correlations overall.

Spatial Distribution of Local Spread-Error Correlation

Domain-Averaged STD-AEM correlation ~ 0.62



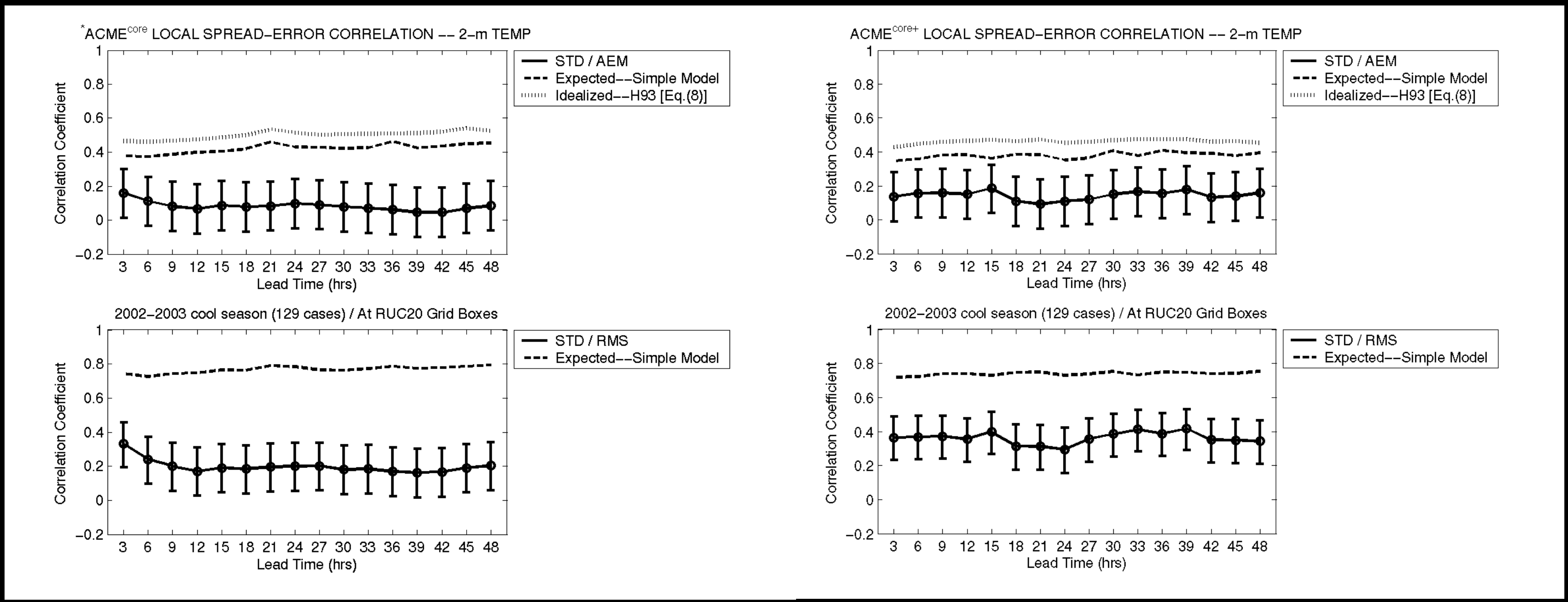
Maximum Local STD-AEM correlation ~ 0.54



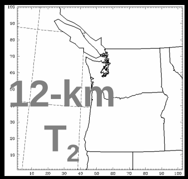
Average Local Spread-Error Correlation

(no bias correction)

ACME^{core} ACME^{core+}



- The average local spread-error correlations are small.
- Estimates from the simple stochastic model are more applicable here, giving an indication of the departure from local statistical consistency.



Preliminary Conclusions

- Accounting for model and surface boundary parameter uncertainty in a mesoscale SREF system is crucial.
  - ACME^{core+} forecasts possess valuable information about the flow-dependent mesoscale uncertainty that ACME^{core} forecasts do not.
- Eckel and Mass (2003) found that a simple bias correction improves ensemble forecast skill, but these results suggest that degradations are also possible.
  - Traditional spread-error correlations are reduced in many cases
  - A shorter range of error variances are resolved (F00-F15)
- Continuous (categorical) predictors of forecast error are most appropriate for end users with a continuous (categorical) utility function.

Outstanding Questions

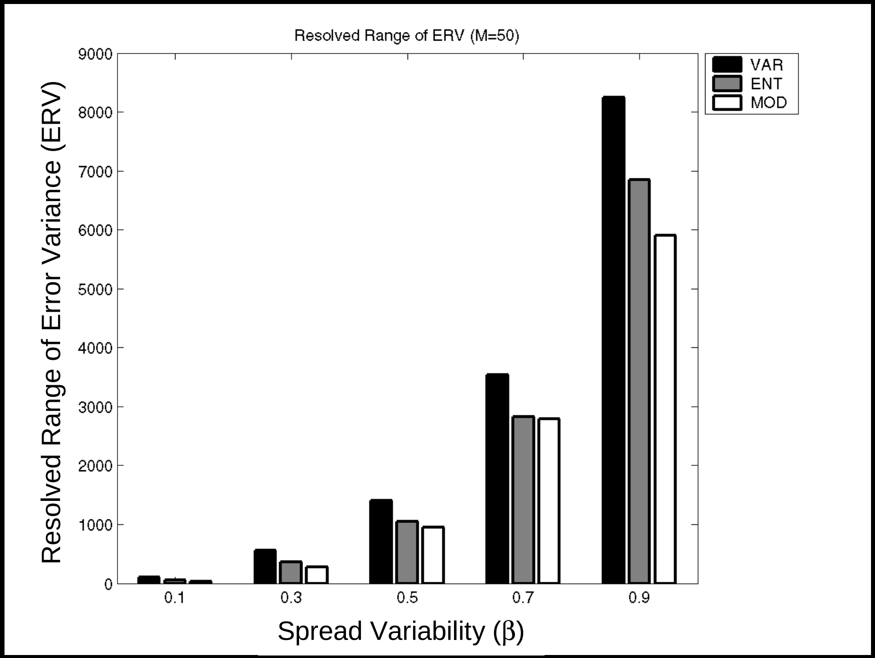
- How can an ensemble-based prediction system for <u>local</u> forecast errors be developed?
  - Ecological (domain-averaged) spread-error correlations can be quite large, while local spread-error correlations are near zero.
  - Can we ever expect local statistical consistency?

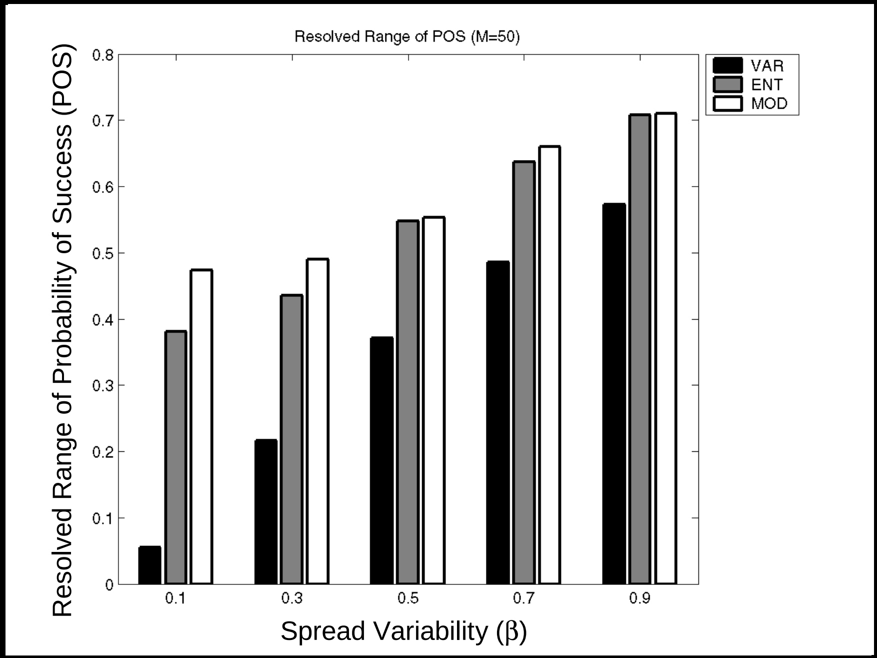
- Will more sophisticated post-processing methods (e.g. – ensemble MOS, best-member dressing, Bayesian model averaging) also degrade the forecast error predictability?
  - Or is the decrease in forecast error predictability in this study an aberration?
  - Maintaining case-to-case spread variability must be a constraint of paramount importance for ensemble post-processing methods.

FURTHER QUESTIONS???

EXTRA SLIDES

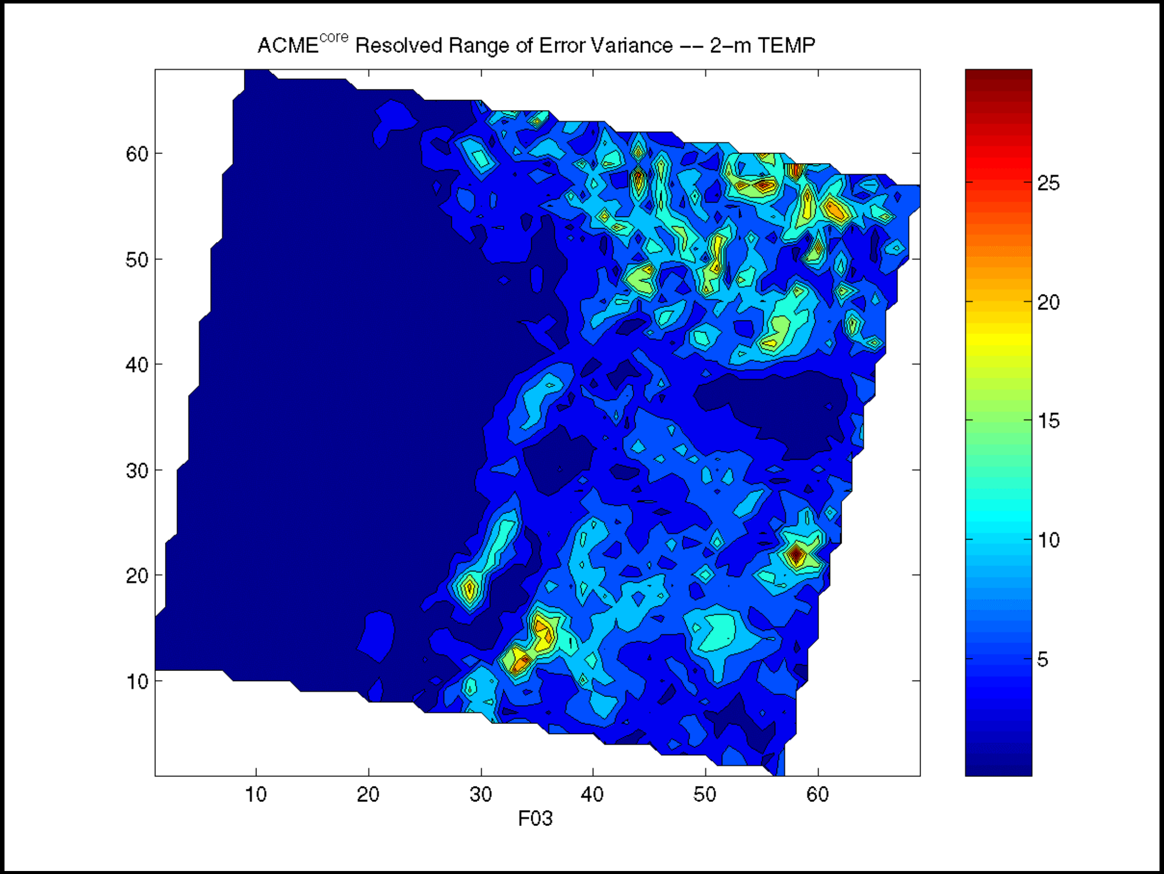
Continuous or Categorical Predictors?

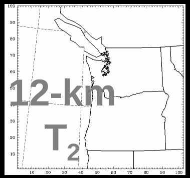




- Continuous (categorical) predictors of forecast error are most skillful for end users with a continuous (categorical) utility function.

Spatial Distribution of Resolved Range of ERV

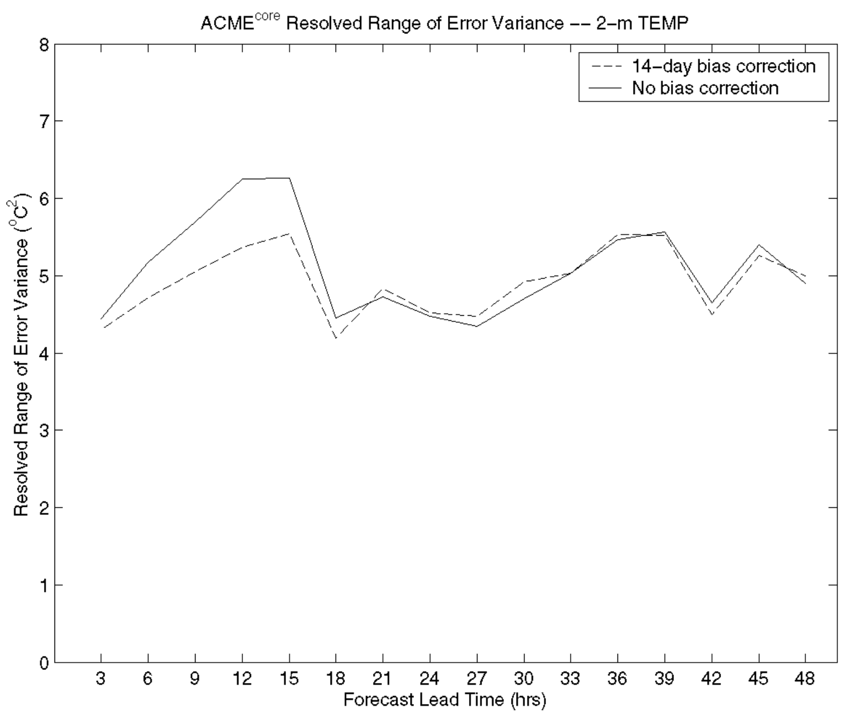


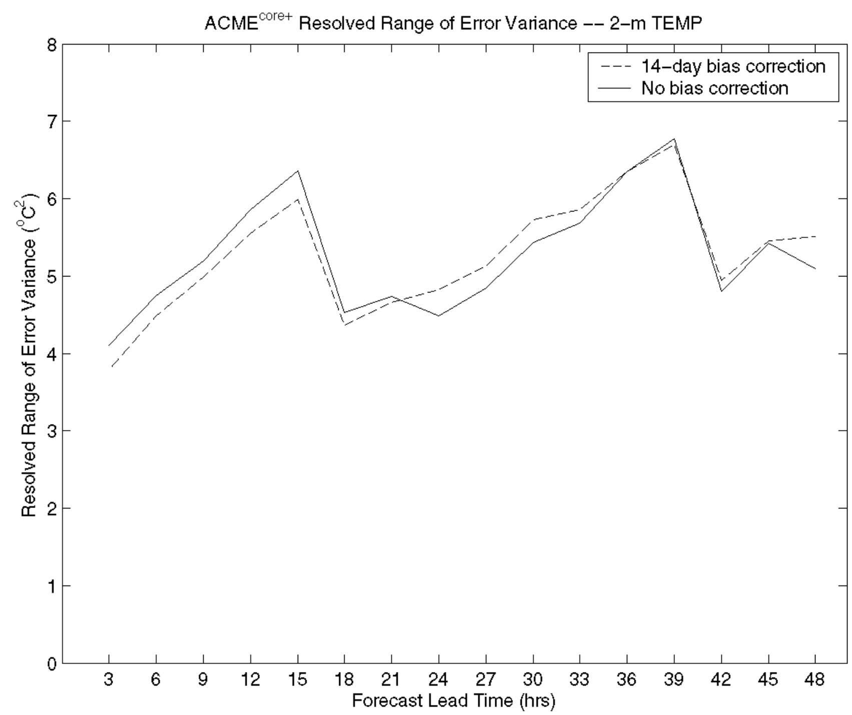


Resolved Range of Local Error Variance

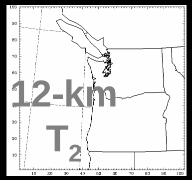
(Domain-averaged)

ACME^{core} ACME^{core+}





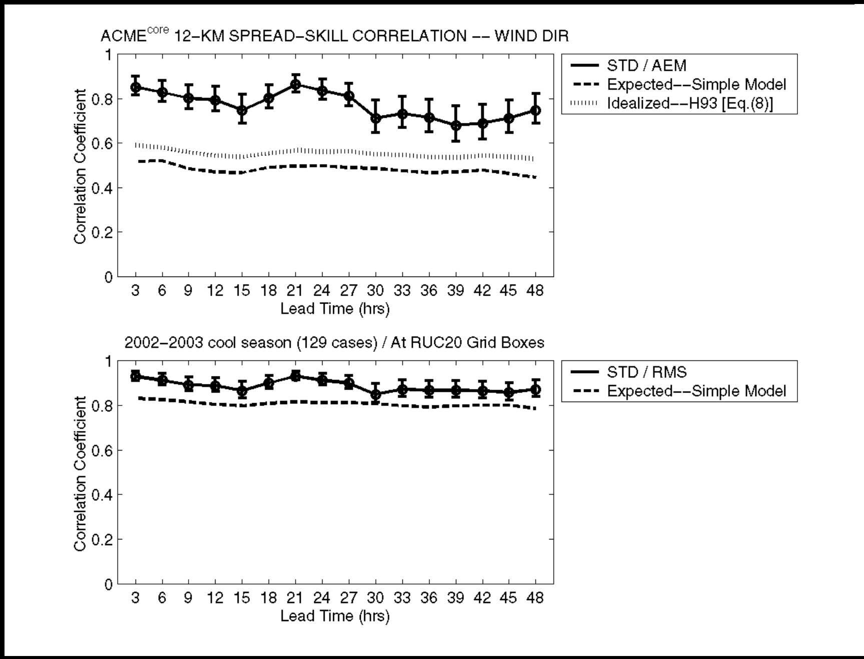
- Bias correction reduces the resolved range of local error variances during the first 15h. At longer lead times, no difference is apparent.



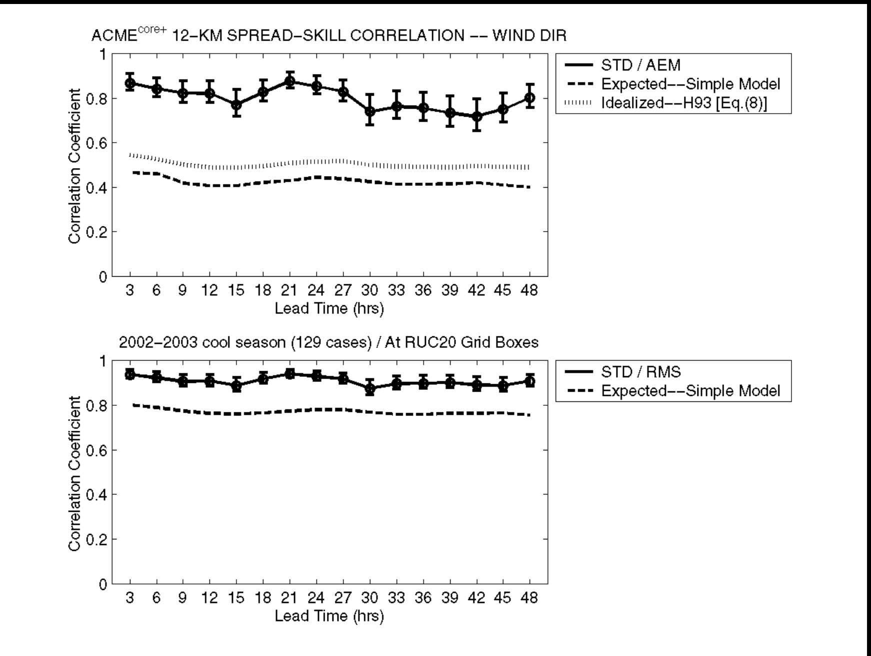
Spread-Error Correlations

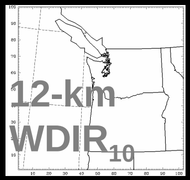
(no bias correction)

ACME^{core}



ACME^{core+}



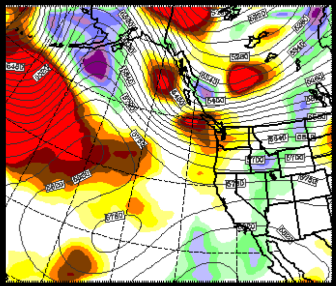


Forecast Error Prediction

- Like any other scientific prediction or measurement, weather forecasts should be accompanied by error bounds, or a statement of uncertainty.

$$T_{2m} = 3\ ^\circ C \pm 2\ ^\circ C \qquad P(T_{2m} < 0\ ^\circ C) = 6.7\ \%$$

- Forecast uncertainty changes spatially and temporally, and is dependent on:
  - Atmospheric predictability – a function of the sensitivity of the flow to:
    - Magnitude/orientation of initial state errors
    - Numerical model errors / deficiencies



- Ensemble weather forecasts appear well-suited for quantifying fluctuations in atmospheric predictability

Value of Forecast Error Prediction

- Operational forecasters require explicit prediction of this flow-dependent forecast uncertainty
  - Helps to decide how much to trust model forecast guidance
  - Current uncertainty knowledge is partial, and largely subjective

- End users could greatly benefit from knowing the expected forecast error
  - Allows sophisticated users to make optimal decisions in the face of uncertainty (economic cost-loss or utility)

    Take protective action if: $P(|E_{T_{2m}}| > 2\ °C) > \text{cost/loss}$

  - Common users of weather forecasts – confidence index